

| Document Name                       |   |   |                         |
|-------------------------------------|---|---|-------------------------|
| <b>Document Title</b>               | InSPECT Framework Report                                      |   |                         |
| <b>Work Package</b>                 | 3.3   |   |                         |
| <b>Author(s) &amp; project role</b> | Gareth Knight, Digital Curation Specialist                    |   |                         |
| <b>Date</b>                         | 13 October 2009   | <b>Filename</b>   | inspect-methodology.pdf |
| <b>Access</b>                       | <input checked="" type="checkbox"/> Project and JISC internal | <input checked="" type="checkbox"/> General dissemination |                         |

| Document History |                |   |
|------------------|----------------|---|
| Version          | Date           | Comments  |
| 2.1              | 13 October 09  | Review and transfer into template   |
| 2.0              | 23 May 09      | Revision of framework report on basis of lessons learnt from case studies and incorporation of FBS framework. |
| 1.0              | 14 February 08 | Revision based on feedback by project team (Paul Wheatley, Lynne Montague & Adrian Brown)                     |
| 0.3              | 5 February 08  | Revision based on feedback by project team (Paul Wheatley, Lynne Montague & Adrian Brown)                     |
| 0.2              | 12 January 08  | Creation of preliminary list of significant properties for audio, structured text and email                   |
| 0.1              | 10 October 07  | First draft   |

# Table of Contents

|  |    |
|--|----|
| Table of Contents .....  | 2  |
| 1. Introduction .....  | 3  |
| 1.1 Introduction to Digital curation .....   | 3  |
| 1.2 Interpreting digital information for access and use.....                               | 3  |
| 1.3 Significant properties as a component of preservation .....                            | 4  |
| 1.4 Significant properties and the OAIS .....  | 5  |
| 1.5 Digital curation strategies .....  | 7  |
| 2. Framework for determining significance.....   | 8  |
| 3.1 Literature review .....  | 8  |
| 3.2 Analysis methodology.....  | 9  |
| 3.3 Assessment Framework .....   | 9  |
| 3.4 Applying the concept of artefact design and management to the curation lifecycle ..... | 10 |
| 3. Requirements Analysis.....  | 13 |
| 3.1 Object analysis .....  | 13 |
| 1. Select object type for analysis.....  | 15 |
| 2. Analyse structure.....  | 15 |
| 3. Identify purpose of technical properties .....  | 16 |
| 4. Determine expected behaviours .....   | 16 |
| 5. Classify behaviours into functions.....   | 17 |
| 6. Associate structure with each behaviour.....  | 19 |
| 7. Review and finalise .....   | 20 |
| 3.2 Stakeholder requirements analysis .....  | 21 |
| 1. Identify stakeholders .....   | 22 |
| 2. Select object type for analysis.....  | 22 |
| 3. Determine actual behaviours.....  | 23 |
| 4. Classify behaviours into set of functions .....   | 24 |
| 5. Cross-match functions.....  | 25 |
| 6. Assign Acceptable value boundaries .....  | 25 |
| 7. Review and finalise .....   | 26 |
| 3.3 Reformulation .....  | 27 |
| Significant Properties Data Dictionary .....   | 30 |
| References .....   | 31 |

# 1. Introduction

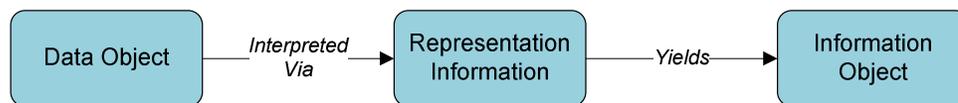
## 1.1 Introduction to Digital curation

Curation refers to the sequence of activities necessary to maintain information over a period of time. The notion originates in the physical realm, where it is commonly used to describe the management of information stored on an analogue carrier. In the 1990s, the notion of digital curation emerged to refer to the process performed to ensure that digital information remains accessible throughout its lifecycle. The curation of physical and digital objects has a common objective - to ensure that information remains accessible in the long-term. However, the method in which they are achieved differs considerably. In the physical realm, curation often refers to the process of maintaining the physical carrier on which information is stored. In the digital realm, it is insufficient to store the source carrier and expect it to be usable in the long-term. Instead, digital curation represents a recognition that technological change is inevitable and that active management is required to maintain access to information stored in a digital form in the long-term.

## 1.2 Interpreting digital information for access and use

To understand the purpose of digital curation and the role of significant properties, it is necessary to examine the method in which digital information is stored and the process by which it is accessed. The OAIS Reference Model indicates that the recreation of information requires two components: (1) a Data Object that contains the information in an un-interpreted form and (2) Representation Information necessary to decode the Data Object and recreate it as an Information Object. The National Archives of Australia offer a similar interpretation, using the simile of a performance to illustrate the process (Heslop, Davis & Wilson, 2002).

OAIS Reference Model



NAA Performance Model

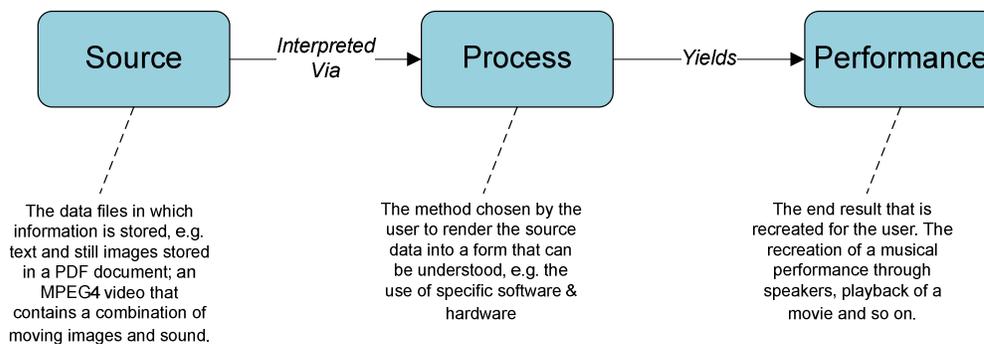


Figure 1: Data interpretation and recreation

The key stage to examine in the Performance model is 'Process', the method by which information is interpreted and rendered into an understandable form for the user. For example, a source may be a play script or poem, which would be learnt by an actor and performed to an audience. In its raw form, digital information is stored as binary data encoded on some type of media. A combination of hardware and software is required to interpret the binary data and render it in a form that can be understood by the user, e.g. an audio recording played through computer speakers or a photograph displayed on screen.

In the long-term, some event may occur that alters the method in which the source is interpreted. In the performance simile, the actor is replaced by a second actor that uses a different interpretation of the source, resulting in a different performance to an audience. A similar event may occur in the digital realm - changes to the computing environment may result in some change in the recreation of the information. For example, the computer hardware, operating system and application software in use in five years may be only semi-compatible with digital objects stored in older formats (Wilson, [2007](#)). The variation in performance introduced through the use of different technology may be relatively minor (e.g. use of an alternative font) or may result in major changes (e.g. content loss or corruption).

### **1.3 Significant properties as a component of preservation**

The Oxford English Dictionary defines 'significant' as an adjective that refers to something that conveys a meaning or has some importance. Further investigation of the two terms provide definitions and examples of their use: a meaning is commonly associated with a purpose, motive, justification, intention, or other implication; while Import refers to something of value or prominence. The exploration of the definition of significance and related terms result in the recognition of five factors:

1. Significance is relativistic, rather than being universal and unchanging;
2. Interpretations of significance will differ dependent upon the intended purpose and the criteria that is applied
3. Meaning may be intrinsic in the construction of an item.
4. Meaning is conveyed through a process of communication from a source
5. Meaning may be interpreted differently by stakeholders, dependent upon their knowledge base, environment in which they operate and other factors<sup>1</sup>.

The interpretation of significance has specific connotations when used within the archival community to refer to digital objects. The concept of significant properties emerged through work performed by the CEDARS project in the late 1990s to describe the elements of a digital object that should be maintained through preservation action. It is built upon the underlying belief that it may be impractical, due to technical issues, cost or other factors to reproduce all elements of an object over time. For example, a new format may not support all features of the original or an emulator may introduce anomalies into the recreation. Therefore, selection criteria should be developed that enables a curator to determine the elements of an object that must be maintained and distinguish them from those that may be abandoned. Since its development, the concept has been described using several different synonyms (essence, essential characteristics, core features, properties of conceptual object and others) and been subject to different, although semi-compatible interpretations. In an early work package for the InSPECT project, Wilson (2007) reviewed related work and proposed a revised definition. Significant properties, Wilson states, refer to:

*The characteristics of digital objects that must be preserved over time in order to ensure the continued accessibility, usability, and meaning of the objects, and their capacity to be accepted as evidence of what they purport to record*

In later stages of the project, the InSPECT project team attempted to ally the interpretation of significant properties more closely with the OAIS Reference Model and define its relationship with Representation Information. This has resulted in the incorporation of the following reference to the OAIS Information Object in the definition.

---

<sup>1</sup> See Knight, G. (2008c) for further discussion of the topic.

*The characteristics of an Information Object that must be maintained over time to ensure its continued access, use, and meaning, and its capacity to be accepted as evidence of what it purports to record.*

The concept of significant properties using both definitions is associated with the notions of authenticity (that a record is what it purports to be) and integrity (that there has not been corruption at the bit-level or deliberate alteration at the semantic level that has caused the original meaning to be lost) (Bearman & Trant, [1998](#); Digital Preservation Testbed, [2003](#); Wilson, [2007](#)). The requirements to maintain the authenticity and integrity of an object are useful in determining the properties that may be considered significant. However, as evidenced in the definition of significance and associated terms, the interpretations of authenticity and integrity may vary in different contexts.

## 1.4 Significant properties and the OAIS

The Open Archival Information Systems (OAIS) Reference Model (ISO 14721) is an international standard that addresses the requirements of maintaining access to information in the long-term. The reference model provides a conceptual framework and common terminology to describe the concepts, processes and systems associated with digital curation. These may be realised by adopting a set of procedures and practices that fulfill the aims and objectives of an OAIS-compliant system. The reference model is separated into two sub-sections: a Functional Model that outlines a set of archival functions (Ingest, Archival Storage, Data Management, Access, and Administration) and processes that must exist to accept, manage and publish information; and an Information Model that indicates the requirements necessary to access the information over time.

The Information Object has a central role within an OAIS, representing the product that must be recreated in order for a user to understand the information content. An Information Object is realised through the combination of a Data Object (the bit-stream) and Representation Information, the information that enables the Data Object to be interpreted and rendered. In an OAIS, management of the Data Object and Representation Information is essential for maintaining access to the Information Object.

The Information Object is transferred into, through and out of an OAIS within Information Packages. The OAIS defines three types of Information Package: [1] a Submission Information Package (SIP) that is obtained on Ingest; [2] an Archival Information Package (AIP) created in a form suitable for preservation; and a Dissemination Information Package (DIP) created for distribution. Each Information Package may contain a different Data Object that will require Representation Information to interpret it. However, the OAIS indicates that the Information Object should remain the same throughout.

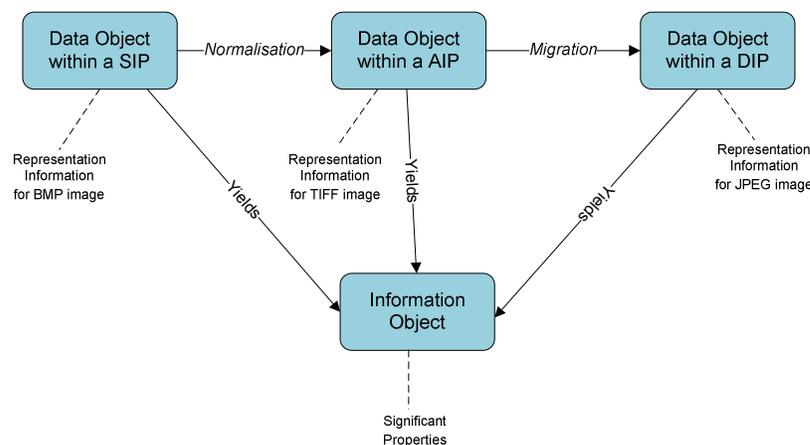


Figure 2: One Information Object, three Data Objects

In the example, Representation Information is necessary to decode a BMP, TIFF and JPEG Data Object and recreate the Information Object in the form of a still image.

An assumption implicit in the OAIS Reference Model is that a single type of Representation Information will exist for each Data Object that will be used to recreate the Information Object. Although ideal, this does not reflect practical experience of accessing a Data Object in a digital environment. As recognised by the National Archives of Australia in its Performance Model, it may be more accurate to recognise the existence of several Representation Information variants for a single Data Object. The use of one Representation Information variant may yield an Information Object that differs from that rendered by a second Representation Information variant. The differences between the two recreations may be considered minor or major dependent upon its influence upon the access, use and interpretation of the Information Object.

To evaluate the accuracy of an Information Object's recreation, or activities performed when transferring it between Data Objects it may be considered useful to record descriptive information about the Information Object itself. The InSPECT project has, through the representation of Brown (2008) in the project team, adopted the position that Significant Properties refer to the aspects of the Information Object that are necessary to support its understanding and use.

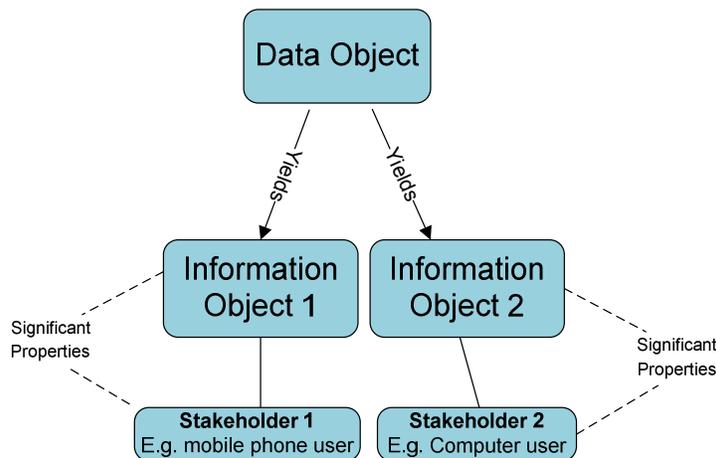


Figure 3: One Data Object, multiple Information Objects

In applying the concept of significant properties to the OAIS Information Object, the InSPECT project recognized the contribution of the Designated Community – “*the subset of Consumers expected to independently understand the archived information*” (Lavoie 2004) – to the interpretation of significance. Rather than maintain a single Information Object that is used by Designated Communities (or stakeholders as described in the project methodology), it is considered likely that variations of the Information Object will be created that contain a subset of attributes from an source Data Object contained in an SIP or AIP. For example, when creating a derivative of an email object for use by mobile phone users, the evaluator may choose to exclude properties associated with the presentational or semantic mark-up of the message that cannot be displayed on some mobile phone or to reduce the size of the Data Object. Alternatively, a small number of the Designated Community may require properties that are not available in the Information Object that is made available as part of a Dissemination Information Package, necessitating the creation of an alternative Information Object for specific stakeholders.

## 1.5 Digital curation strategies

To ensure that the Information Object remains accessible in the long-term term it may be necessary to intervene at specific stages of its lifecycle. Thibodeau (2002) outlines several strategies that an institution may adopt. These include:

1. **Bitstream management:** Management of the original data as a sequence of bits.
2. *Technology preservation:* Maintain hardware and software necessary to access the information in its original form;
3. *Technology recreation:* Reproduce the behaviour of hardware and/or software in a different technological environment (also referred to as emulation or virtual machines);
4. *Format conversion:* Maintain access to information on contemporary hardware and software by converting it from its original bit sequence to an alternative encoding format that may be managed and/or accessed more easily. Normalisation refers to a process of converting something into a 'normal' form, conforming to specific rules or regulations. Migration, for the purpose of this report refers to the process of converting information content into a format that is accessible to the majority of users.

Several projects have been funded over the years, from CEDARS and CAMILEON projects in the late 1990s to the current PLANETS and CASPAR projects with the task of developing and testing management strategies based upon each approach. As a result of the activity, many institutions have adopted management strategies that use a combination of bit-stream management and format conversion or technology recreation. Although pragmatic, the latter two strategies cannot be performed without risk that content will be unexpectedly changed or corrupted. The change may be relatively small – the removal of transparency from an image when converting between GIF and TIFF, for instance – but have major implications for the meaning that is communicated to the user. By recording the significant properties of an Information Object, the curator can evaluate the success of its recreation in a different environment or conversion into a different format and identify aspects that have not been reproduced correctly.

## 2. Framework for determining significance

A formal framework is required to guide the process of identifying, analysing and recording the elements of an Information Object that are essential or beneficial to maintain over subsequent manifestations of a digital object. The framework should be rational in order to support the decision-making process, consistent in its application, while offering sufficient flexibility to meet the needs of the evaluator. By applying the framework, an evaluator should be able to make an informed choice based upon consideration of associated factors, rather than haphazard decisions that cannot be supported at a later date. The following sections outline the investigative process performed and the framework that was developed by the InSPECT project to determine the set of requirements for raster images, audio recordings, structured text and e-mail.

### 3.1 Literature review

The importance and position of significant properties in the development of digital preservation strategies have been recognised by several parties over the last decade and, as a result, there has been a broad body of work considering how they may be identified and analysed. In the first year of the project, the project team reviewed a broad range of literature written on the topic to identify if any had developed frameworks that may be used for identification and evaluation (Knight, 2008; Wilson, 2007). Notable work that was examined includes that written by Rothenberg & Bikson (1999), the CEDARS Project, the CAMiLEON project, the National Archives of Australia, RLG, Digital Preservation Testbed, DELOS, as well as more recent developments by the CASPAR, PLANETS and four JISC-funded Significant Properties projects. Many of these studies describe the process by which they identified the significant properties of various object types, describing formalised and semi-formal methods.

The more formal frameworks outline a set of activities that an evaluator should perform to obtain a list of properties that are necessary for preservation. In an early study of the topic, Rothenberg & Bikson developed a Needs Analysis model and identify four keys stages<sup>2</sup> that should be followed to determine the elements of an Object that must and can be maintained. This was followed by the InterPARES1 project, which applied the principles of archival diplomatics to digital records as a method of determining its authenticity. The methodology is built upon the premise that many of the authenticity requirements of a record can only be determined by considering its intended purpose in an organizational setting and, as a result, cannot be easily understood by examining the record in isolation. However, the InterPARES Authenticity Task Force recognized that archival diplomatics as a methodology is tailored to requirements that may be identified in a known organizational environment and, as a result, it is difficult to apply them to digital records that do not contain textual information, are dynamically produced, or published in a system where the context cannot be known (MacNeil et al, 2000). Finally, the conceptual Utility Analysis and Objective tree (Rauch, Strodl & Rauber, 2005) was applied in the DELOS and PLANETS project as a metric to test and evaluate digital preservation strategies. A key development is the specification of four main groups of characteristics: object, record, process and costs as a basis for evaluating different preservation options.

The frameworks developed by each project were useful in informing the development of the project, but were considered to be insufficient in isolation. In developing its methodology, the InSPECT project drew upon elements expressed in several projects. The needs analysis approach in the RAND-Europe study was recognised as being particularly useful for determining the requirements of stakeholders within specific environments<sup>3</sup>. The expression of specific

---

<sup>2</sup> The stages are: 1) Analyse the functions that the records must support; 2) define authenticity criteria; 3) decide the records to be preserved; and 4) Analyse technological alternatives for preservation that result in a final stage in which an appropriate preservation strategy is chosen.

<sup>3</sup> Evidence of the approach can be found in the CEDARS project, which used group discussion as a method of gathering information for identifying significant properties.

requirements as an Objective Tree by the PLANETS project was also seen as an interesting approach, but at the time it was evaluated, it was recognised that a different approach may be necessary that placed a greater emphasis upon the role that properties performed for the recreation of the Information Object.

### **3.2 Analysis methodology**

The methodology adopted to determine the properties to be maintained in subsequent manifestations of a digital object was developed on an ongoing basis during the lifetime of the project. Experience gained by the InSPECT project team and the four JISC-funded studies when examining the significant properties of different object types were noted and incorporated into subsequent versions of the analysis methodology. The project also drew upon feedback provided by email and through various workshops

It was recognised early in the project that it is impractical to present a single, definitive interpretation of significance. Instead, it should adopt a methodology that enables the evaluator to identify the stakeholders that have some investment in the Information Object and define the subjective decision-making process that contributes to their evaluation of significance. Key to this approach was the recognition of three factors:

- 1) Many stakeholders may be associated with an Object, e.g. creator, researcher;
- 2) The type of stakeholder associated with the Object may vary and change at different stages of its lifecycle<sup>4</sup>. For example, a creator may use an Information Object in the initial stages of its lifecycle and subsequently make it available for use by other researchers.
- 3) Each stakeholder may possess a distinct knowledge base and have specific needs for the task they wish to perform.

By adopting a relativistic approach, an evaluator operating in a curatorial institution can determine the properties that they consider to be essential based upon their interpretation of acceptable loss. They may accept that some loss of functionality is necessary if it is to simplify the preservation process or, alternatively may adopt a risk adverse approach and adopt a preservation strategy that enables them to maintain all properties of the Information Object.

The methodology was underpinned by a joint teleological and epistemological approach. Teleology is the philosophical study of the design and purpose of an object. This conceptualises an author as a designer that creates an object as the result of an intellectual process to fulfill specific objectives or to address a problem. Epistemology is a branch of philosophy that is concerned with the meaning of knowledge and the process by which knowledge is acquired. In combination, the two philosophical branches require the evaluator to determine the context of the object's creation (the purpose it was created for, how it was created, and so on) and the information necessary to communicate the intrinsic knowledge to a designated community.

### **3.3 Assessment Framework**

The first version of the assessment framework (Knight, 2008b) outlined a set of activities that an evaluator could follow and offered a template to record values. To follow the instructions, an evaluator would start by examining the Information Object in its entirety (e.g. an email, raster image, audio recording) and progress through the sub-components until they can identify the technical properties that are necessary to recreate it. Once the evaluator has reached the

---

<sup>4</sup> Many types of lifecycle may be associated with a digital object. For example, a digital lifecycle that is influenced by the ability to access content and an Information lifecycle that refers to stages of use within an environment.

property-level, they would work with one or more stakeholders in the Designated Community to analyse the acceptable boundaries necessary to achieve their stated objectives.

The current version of the assessment framework utilizes design methods to identify and evaluate the functions performed by an Object in its current manifestation and re-develop it to meet the needs of other stakeholders, such as a curator. To structure the analysis process, the project adopted a modified version of the Function-Behaviour-Structure (FBS) framework. The framework was initially developed by John Gero in 1990 to assist engineers and designers with the process of creating and re-engineering systems and has been revised and refined on several occasions since. Gero interprets the design process as an intentional, intellectual process by which a designer takes a set of designated functions and transforms them into a design description for an artefact structure that can fulfill these requirements. In the FBS model, the behaviour that is exhibited by an artefact (e.g. the operation of a motor engine) is a product of the functional purpose established by the designer and the physical structure which compose the artefact. The design method may also be used to reverse-engineer and re-design an existing product to perform one or more new functions (Takeda et al, n.d). For example, the structure of a bicycle may be re-designed to enable it to fit into a small space for transport when not in use.

Although the role of an engineer that is responsible for re-designing an artefact may initially appear to conflict with the curatorial duty to maintain the Information Object, they are not so dissimilar. As Rusbridge (2006) notes, it may not be necessary to be faithful to the original object in all respects. Many stakeholders may be willing to accept an Object that omits specific content or functionality. By considering the purpose of an object in conjunction with the stakeholder that uses it, a curator may identify the functions required by the creator in the early stages of the object lifecycle and evaluate if they continue to be necessary when used by a different community of users. This may result in the curator recognizing that some functionality is not required and adopting a preservation strategy that is faster, simpler to perform and less costly than alternative strategies that maintain all elements of the object<sup>5</sup>.

In applying the FBS model to the analysis of digital objects, the InSPECT project has reinterpreted the base terminology used and the set of activities necessary to perform the requirements gathering process. The following definitions are used of the titular components of the FBS framework:

1. *Function*: The design intention or purpose that is performed.
2. *Behaviour*: The epistemological outcome derived from the function and structure that is obtained by the stakeholder. E.g. an interpretation of the meaning contained in the Content Information.
3. *Structure*: The structural elements of the Object that enables the stakeholder to achieve the stated behaviour.

The interpretation of Behaviour differs from the definition initially provided by Gero, which examines the behaviour that an artefact exhibit. Instead, an interpretation has been made similar to that provided by Stalker (2002) when examining the lifecycle of an artefact.

### ***3.4 Applying the concept of artefact design and management to the curation lifecycle***

The design process defined by Gero and his collaborators establishes eight steps that a designer may follow to transform a set of proposed function into the design description for an artefact (table 1). Steps 1-5 of the FBS model are intended to be followed sequentially to transform the

---

<sup>5</sup> For example, an institution that uses format conversion as a preservation strategy may choose to export database tables stored in Microsoft Access to tab-delimited format, which would result in loss of the ability to manipulate table information

set of functions that must be performed into a design document for production. If the designer is dissatisfied with aspects of the prototype they may repeat an earlier step until they are satisfied with the design documentation. Steps 6-8 refer to activities that the designer performs to re-design the structure to perform the same functions, provide new functionality, or offer new expected behaviours (Dorst & Vermaas, 2005).

|                 |  |
|-----------------|--|
| Formulation     | Designer takes a set of defined functions ( $F$ ) and derives a set of expected behaviours ( $Be$ ) for the artefact to be created |
| Synthesis       | Designer develops a specification for the structure that will exhibit the 'expected behaviours ( $Be$ ).                           |
| Analysis        | Following the creation of a prototype, the designer analyses the actual behaviours of the structure                                |
| Evaluation      | The designer evaluates the suitability of the actual behaviour by comparing it to the expected behaviours                          |
| Documentation   | A design description is produced for production  |
| Reformulation 1 | Designer chooses a new structure to exhibit behaviours   |
| Reformulation 2 | Designers defines a new set of expected behaviours   |
| Reformulation 3 | Designer chooses new functions to be performed.  |

Table 1: Steps of the design prototype model

Further work by Stalker (2002) has extended the FBS framework to describe the design management of an artefact throughout its lifecycle. In the revised model, steps 1-5 specified in table 1 are followed by a construction stage in which the artefact is created. This may be followed by subsequent steps, in which the designer monitors the artefact to identify behaviours that may emerge<sup>6</sup>, intervene to add new functionality, modify or repair the artefact to allow it to better perform an existing function<sup>7</sup> and possibly dismantle it when it no longer serves a purpose. These additional steps are outlined in table 2.

|                                  |   |
|----------------------------------|---|
| Monitoring                       | The design engineer monitors the artefact to observe new 'actual behaviours' that emerge from its use.  |
| Intervention                     | As a result of the monitoring activity, a design engineer may intervene to correct issues that require a re-design or reformulation of functions. |
| Retrofit / modification / repair | The design engineer repairs or modify the structure of an artefact to provide the required functionality.   |
| Dismantling                      | An artefact that is no longer required or is not fit for purpose is dismantled.   |

Table 2: An extension to the FBS design prototype proposed by Stalker

The design method serves as a useful metaphor for understanding the decisions made and activities performed through the lifecycle of a digital object. These similarities are illustrated in figure 4, which maps the FBS model work performed by Gero (1990) and Stalker (2002) onto the DCC curation lifecycle model.

<sup>6</sup> She cites the unexpected popularity of SMS text messaging as an example of an emergent behaviour.

<sup>7</sup> For example, the Millennium Bridge in London was modified to correct a wobble that occurred when it was used.

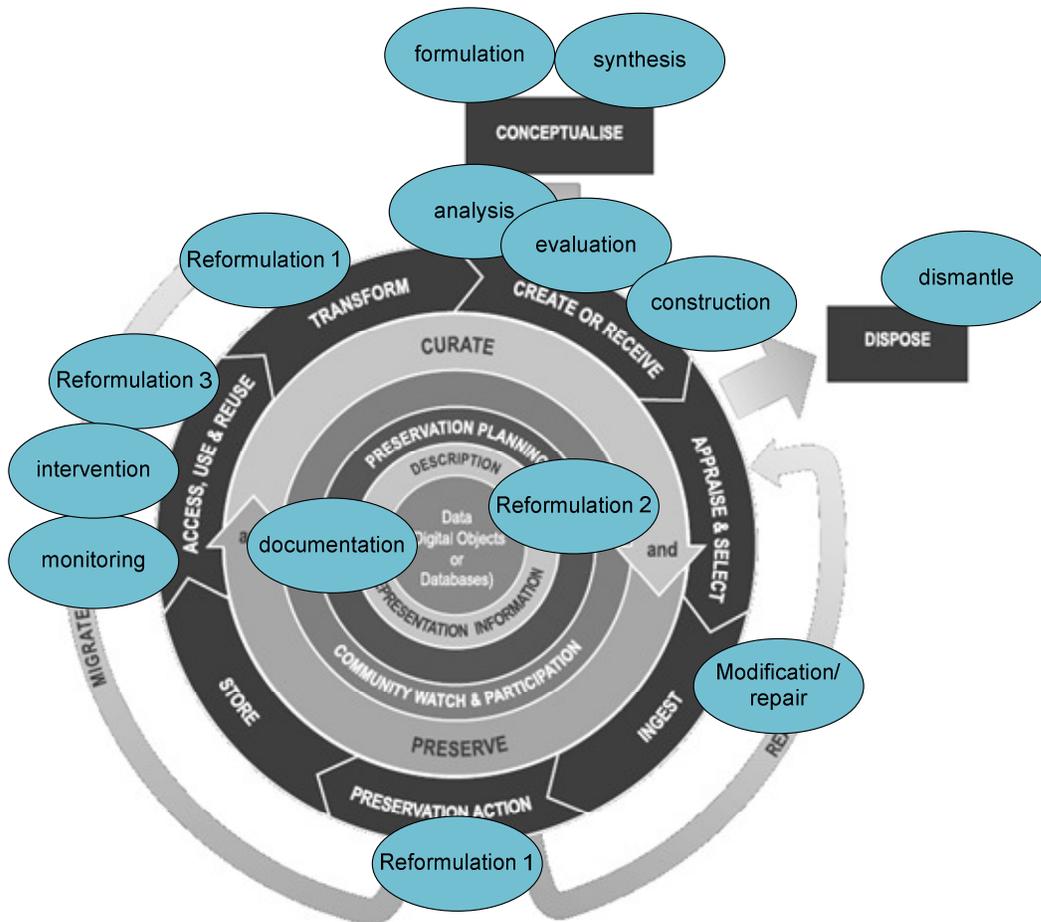


Figure 4: FBS design steps mapped onto the DCC curation lifecycle model

The initial steps of Formulations and Synthesis may be mapped to the DCC Conceptualise stage, in which the creator conceives and plans the creation of data. During the creation process they may analyse and evaluate a prototype and revise it accordingly to fit their needs. At some stage when the creator is satisfied with the prototype it is 'constructed' into a final version of the digital object. The creation of a design document, as expressed in the FBS model may not be present with the creation process of many types of object. However, by re-interpreting the activity as the documentation of the final iteration of the object, design documentation may be used to refer to the creation of Representation Information.

A curatorial institution may also make decisions at a later data that involve some type of reformulation of the object. For example, a digital repository may perform corrective actions at Ingest to resolve issues found in the received object, e.g. modify the encoding structure to conform to a specification or standard) (modification/repair). In Preservation Action/Transform a new encoding structure may be chosen to represent the set of behaviours of the original (reformulation 1). Subsequently, it may be necessary to revise existing behaviours or define new ones that the preservation manifestation of the Information Object should demonstrate (Reformulation 2). Finally, it may be recognized over time that the Designated Community use the object differently than that envisaged and intervention is necessary to offer new functionality or to re-design the structure (Reformulation 3)<sup>8</sup>.

<sup>8</sup> An exploration of the actions necessary to re-formulate an Information Object is outlined on page 27.

### 3. Requirements Analysis

The assessment framework developed by the InSPECT project utilizes the FBS design method to identify the functions that have been defined by the creator of a digital object and evaluate if it is necessary to recreate them in subsequent manifestations of the Information Object. It may also be used to identify if new functions are required to fulfill the needs of stakeholders in the Designated Community<sup>9</sup>. The workflow is composed of three sets of activities (each composed of several sub-tasks): Object Analysis, Stakeholder Analysis and Reformulation.

In Requirements Analysis the evaluator is required to gather information on the existing functionality that is provided by a digital object and understand the tasks that a stakeholder wishes to perform. The activities outlined in this stage may be classified in the 'Define Requirements' stage of the PLANETS preservation planning workflow (Rauch, Strodl & Rauber, 2005). The information gathered during the stage may be subsequently used to determine the properties of the Data Object that are significant for the recreation of an Information Object (e.g. as part of an OAIS AIP or DIP), that is accessed and used by stakeholders in a particular environment. Requirements Analysis is composed of two streams of activity that each possesses a sequential set of sub-tasks to be performed:

- [1] Object Analysis: The evaluator analyses a representative sample of an object type, identifies a set of functions and behaviours that may be achieved, and the properties that are necessary for their performance.
- [2] Stakeholder Analysis: The evaluator identifies one or more stakeholders that have some relationship with the Information Object and analyse the functions that they wish to perform.

The two streams of activities may be performed in parallel or at different time periods. The latter is recommended, to enable the evaluator to gain a greater understanding of the functionality that is provided by the object type, which may be used as a basis for understanding the functions that may practically be provided to a stakeholder.

#### 3.1 Object analysis

In the Object Analysis stage the evaluator selects an Object type for examination and develops their understanding of its technical composition and the purpose for which it may be used. The Object analysis workflow is composed of seven sub-tasks that may be performed sequentially (figure 5).

---

<sup>9</sup> The latter was considered outside the scope of the work performed in the InSPECT project, but may form the basis for future investigation.

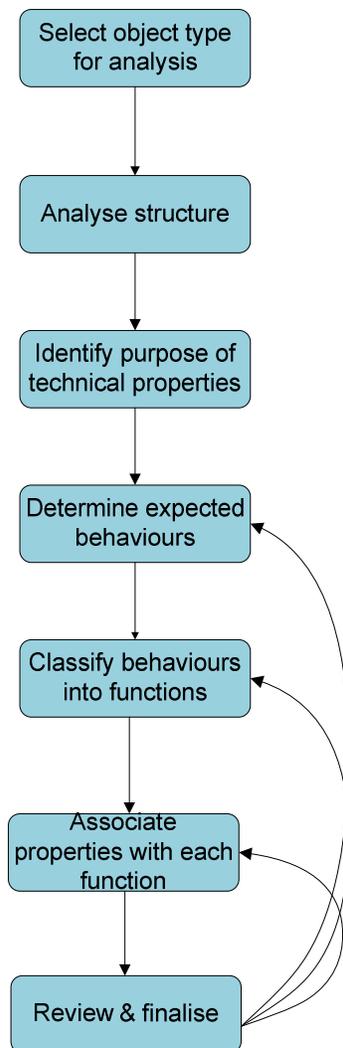


Figure 5: Workflow for the object analysis stage

### Requirements

To perform this stage the evaluator must possess the following:

- A representative sample of objects for analysis
- Technical specifications or standards that describe the composition of the object
- Characterisation tools for analysis of the objects

## 1. Select object type for analysis

The first step is to select the object type to be analysed. The evaluator may choose to select a high-level object type (raster images, audio recordings, web pages, e-mail) or a sub-type that contains specific characteristics.

### Example

An object may be decomposed into sub-types on the basis of different criteria. Sub-types may include:

#### *E-mail sub-types:*

| Criteria                       | Examples  |
|--------------------------------|---|
| Representation format          | HTML email; Rich Text email                                 |
| Relationship with other emails | Standalone email; email message within a discussion thread; |
| creator-defined sub-types      | Artistic email; scientific email                            |
| Institutional sub-type         | Work email; leisure email                                   |

#### *Sound recording sub-types*

| Criteria            | Examples   |
|---------------------|--|
| Audio content       | Vocal recording; music recording   |
| Data object content | Audio recording only; audio recording with descriptive metadata; audio recording with lyrics |

## 2. Analyse structure

Second, the evaluator should analyse the object and obtain a complete list of technical properties. The objective of the task is to develop an understanding of the type of technical properties and value types that are contained within the object type. Each property will be analysed in further detail in the following step.

The task may be performed using several methods. For example,

1. A characterization tool may be used to analyse and extract information on the technical composition of the object for storage as Representation Information.
2. The evaluator may review technical specifications or standards associated with the object type and identify the technical information that is used to construct the Data Object.

### Example

Technical properties for various object types may be found in the following reports produced by the InSPECT project:

- Significant Properties Testing Report: Raster images
- Significant Properties Testing Report: Digital audio recordings
- Significant Properties Testing Report: Presentational mark-up
- Significant Properties Testing Report: Electronic Mail

### 3. Identify purpose of technical properties

Third, the evaluator should determine the purpose of each technical property that composes the object type/sub-type. The purpose of the activity is to determine the role that the property performs within the Data Object. If the technical property contributes to the recreation of the Information Object, it is considered useful to record the property value, for later evaluation after preservation action has been performed.

When analysing technical properties that may be associated with the Information Object of raster images, audio recordings, presentational mark-up and e-mail, the InSPECT project used the following categories:

1. *Content*: Information content within the Information Object. For example, text, still and moving images, audio, and other intellectual productions. Examples: duration, character count.
2. *Context*: Any information that describes the environment in which the Content was created or that affect its intended meaning. Examples: Creator name, date of creation.
3. *Rendering*: Any information that contributes to the re-creation of the performance. For example, font type, colour and size, bit depth.
4. *Structure*: Information that describes the extrinsic or intrinsic relationship between two or more types of content, as required to reconstruct the performance. E.g. e-mail attachments.
5. *Behaviour*: Properties that indicate the method in which content interacts with other stimuli. For example, hyperlinks<sup>10</sup>.

The five terms may be used as high-level categories to distinguish properties of the Information Object from those of the Data Object. Each term may be further decomposed into sub-elements. E.g. Context: provenance, context: descriptive

#### Example

Examples may be found in the following reports produced by the InSPECT project:

- Significant Properties Testing Report: Raster images
- Significant Properties Testing Report: Digital audio recordings
- Significant Properties Testing Report: Presentational mark-up
- Significant Properties Testing Report: Electronic Mail

Other examples may also be found in the XCDL/XCEL work performed as part of the PLANETS project.

### 4. Determine expected behaviours

Fourth, the evaluator should consider the different types of activities that a user – any type of user – may wish to perform. The list of activities should be recorded as a set of expected behaviours.

At this stage of analysis, the evaluator should consider all uses of the Object type rather than those limited to a particular stakeholder. To produce a list of expected behaviours, the evaluator may draw upon their own experiences, the list of property descriptions performed in the previous

---

<sup>10</sup> Behaviour in the category list is used in the context established by Rothenberg & Bikson.

step, formal standards and specifications, or other information sources. It may also be beneficial to consider the purpose for which the Information Object was utilized in its original creation environment.

#### **Example**

##### Email

A brainstorm of activities that a stakeholder may potentially wish to perform when accessing an email Information Object includes:

- Establish the email account from which a message originated.
- Establish the name of the person who sent the message.
- Establish the placement of the email message within an ongoing discussion.
- View and interpret the textual content of the message
- View and interpret the relational structure of text and images within the message
- Reproduce the visual appearance of the message as defined by the creator.
- Establish the authenticity of the message by examining the route it took to reach the recipient.

## **5. Classify behaviours into functions**

Fifth, the evaluator should classify the set of behaviours identified in the previous stage into a set of functions. The functions may be used as a basis for tailoring future manifestations of the Information Object to the needs of the stakeholder.

In performing the activity, the evaluator may recognize that two or more behaviours may be associated with a single function. Alternatively, they may recognize that other behaviours emerge that should be recorded. For example, the recreation of the visual appearance of a message body may result in the recipient understanding contextual information that is implicit in the visual layout.

## Example

Email

### Functions

Recreate message content

Recreate visual appearance of message

Define Message relationships

Establish message context

Verify authenticity

Identify Sender

Identify Recipients

### Behaviours

View author created text

Understand organisational structure of message

Understand relational structure of text and images within the message

Determine placement of message in ongoing discussion

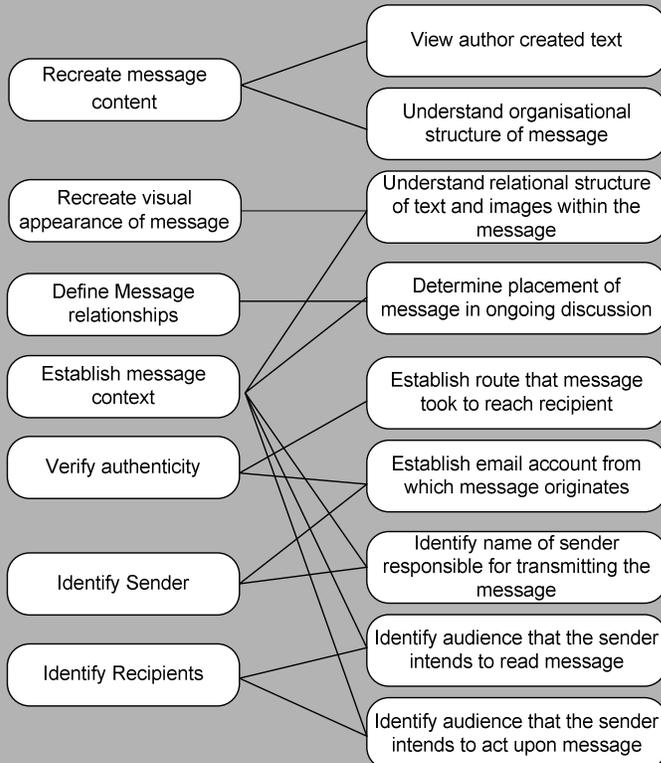
Establish route that message took to reach recipient

Establish email account from which message originates

Identify name of sender responsible for transmitting the message

Identify audience that the sender intends to read message

Identify audience that the sender intends to act upon message



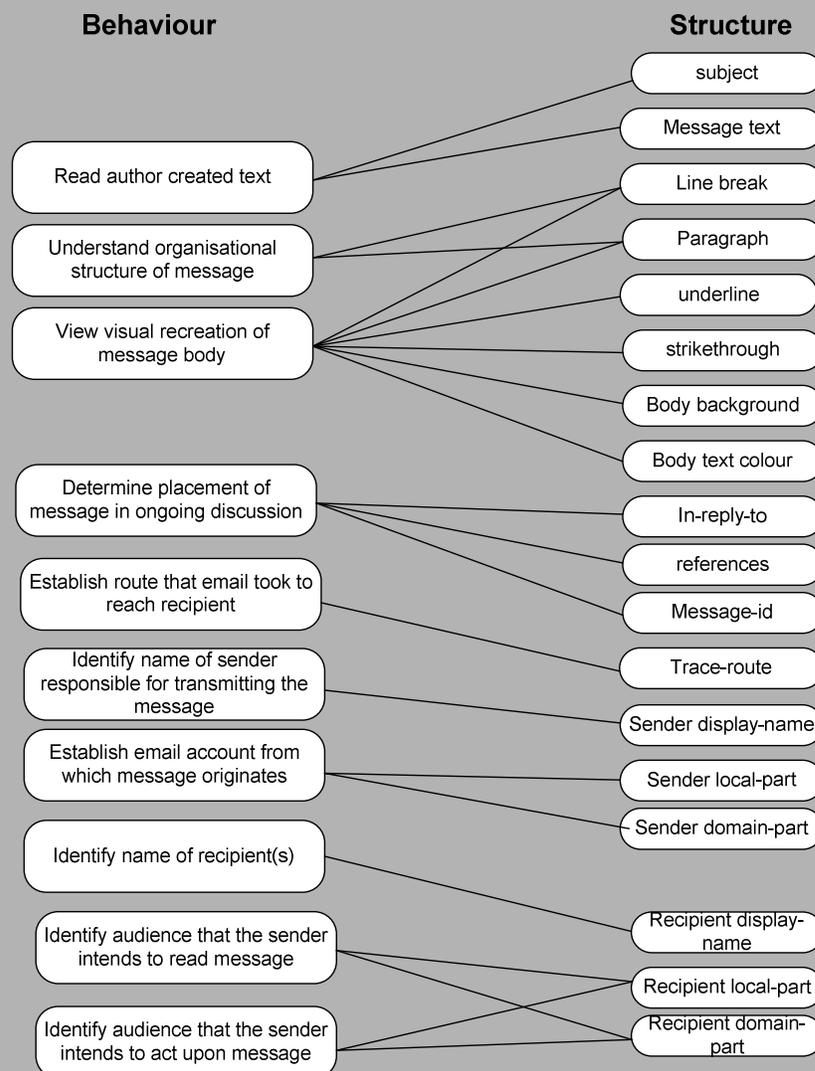
## 6. Associate structure with each behaviour

The purpose of the sixth step is to link the technical properties that establish the structure of the Data Object with the set of expected behaviours. By performing the task, the evaluator may identify and list the subset of technical properties found within the Data Object that contribute to the recreation of the Information Object. The subset of Information Object properties may subsequently be measured and validated when performing format normalization, format migration, or other types of preservation action.

### Example

Email

A set of significant properties associated with an email Information Object may include the following:



The above list is incomplete and provided for illustration only. Please consult the InSPECT Significant Properties Testing Report for further information.

## **7. Review and finalise**

Finally, the evaluator should review the information gathered in the previous steps and consider if any revisions should be made. Pertinent questions to be asked at this stage include:

1. Are there any other behaviours that may be exhibited?
2. Can any of the Functions identified be de-composed into two or more Functions that are more accurate?
3. Are there any other properties that should be associated with a Function?

Once the evaluator is satisfied that they have completed the task, they may record the information accordingly.

### 3.2 Stakeholder requirements analysis

The objective of the stakeholder requirements analysis is to identify the stakeholder categories that may have some relationship with the object type/sub-type and determine the set of functions that they require when using it. The set of functions associated with the stakeholder may be subsequently cross-matched with the object type functions and a list of significant properties developed for each context. When performing the analysis of raster images, structured text, digital audio recordings and e-mail, the InSPECT project team examined the requirements of a curatorial institution. However, the stakeholder requirements analysis may be performed on other stakeholders, such as a creator, researcher, as required by the evaluator.

The workflow for the stakeholder requirements analysis is composed of seven steps. Although they are presented in a sequential order, the evaluator may choose to return to earlier steps at any time to revise it (figure 6).

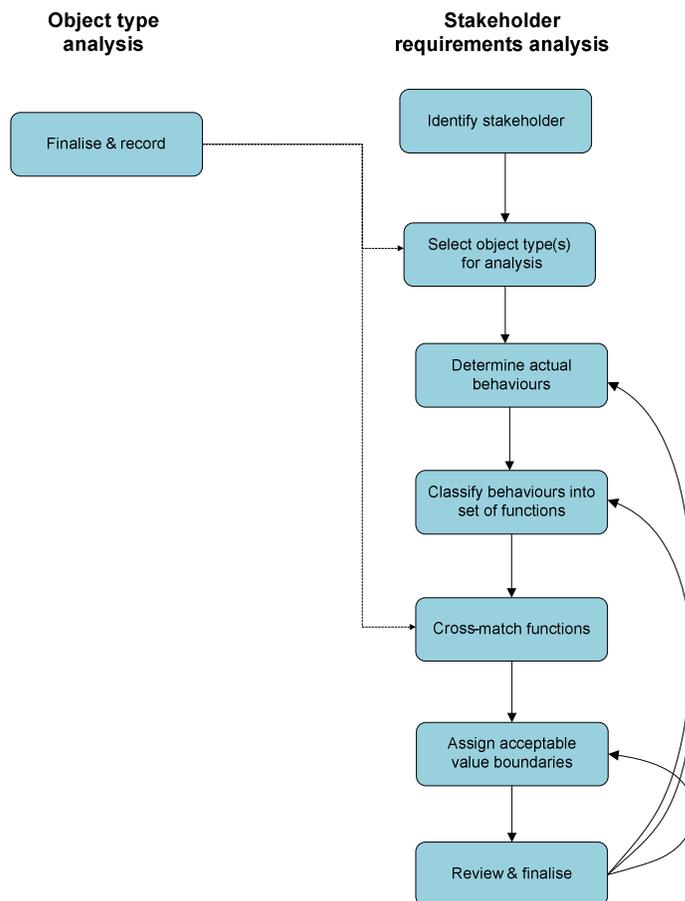


Figure 6: Workflow for the stakeholder requirements analysis

#### Requirements

To perform the analysis the evaluator must possess the following:

- A clear understanding of the relationship between the stakeholder that is the target of analysis and the object type (e.g. researcher, creator, curator)
- One or more people that have been identified as representatives of the stakeholder category

## 1. Identify stakeholders

The first step is to determine the stakeholders that will be the target of analysis and obtain their co-operation. A digital object may be associated with several types of stakeholder throughout its lifecycle, each of which will have different aims and objectives. To identify potential stakeholders for analysis, the evaluator may wish to consult policies, procedures, or legal documents that establish the community that they are intended to serve.

Several methods may be used to obtain details of the actions that a stakeholder will perform when using the object type. Examples may include the use of questionnaires, unstructured/semi-structured/structured interviews and/or observational study. Each research method has benefits and issues that should be considered prior to performing the investigation<sup>11</sup>.

The assessment of a large number of stakeholders may be time-consuming to perform. It may therefore be useful to establish boundaries upon the community that will be examined. The InSPECT project limited the stakeholder analysis to specific curatorial institutions that wish to maintain the authenticity and integrity of a digital object.

### Example

Possible examples of a stakeholder:

- A specific institution that has a defined set of objectives (e.g. a curatorial institution such as The National Archives)
- A role that is working in a particular environment or using specific software tools
- A role that has a direct relationship to the Information Object (e.g. a creator)
- Researchers who wish to analyse content contained in the Information Object using quantitative methods;
- Students who wish to read information, but do not necessarily wish to edit or analyse the data.

## 2. Select object type for analysis

The step is concerned with the selection of an appropriate object type that is used by a stakeholder. As noted in step 1 of the Object type analysis, an object may be classified into a single high-level type (e.g. raster images, audio recordings, web pages, e-mail) or a sub-type that contains specific characteristics. When interviewing the stakeholder, the evaluator may choose to examine the functions required of a specific high-level object type (e.g. raster images, email) or several different sub-types (raster images for scientific use).

It is advised that the evaluator select an object type that has previously been the target of analysis – the list of common functions identified in step 5 of the Object analysis are likely to prove a useful starting point for understanding the functions that a stakeholder may reasonably perform when using a specific instance of the object type.

### Example

Possible examples of a object type/sub-type include:

- Standalone email
- Email messages within a discussion thread
- Emails for use in a work context
- Vocal recordings
- Audio recordings with lyrics
- Audio recordings with descriptive metadata

<sup>11</sup> An evaluation of different research methods for assessing stakeholders in an institution may be found in the Data Audit Framework.

### 3. Determine actual behaviours

The objective of the third step is to determine the activities that a specific category of stakeholder will likely perform when using the object. In the FBS model, the actions that occur in a real-world environment are referred to as 'actual behaviours'<sup>12</sup> and are distinct from the 'expected behaviours' that were defined in step 4 of the object analysis. The actual behaviours may represent a subset of the expected behaviours that were identified (e.g. some users will need to view the Information Object, but may not wish to manipulate or edit the content), or may include new behaviours that were not previously recognized<sup>13</sup>.

To determine the behaviours exhibited by a stakeholder, the evaluator may wish to adopt an epistemological approach by considering the knowledge base that the stakeholder will draw upon and the method in which the content will be interpreted. The set of previously defined expected behaviours may be used to guide a semi-structured interview. Alternatively, they may ask the stakeholder to demonstrate how they would use the Information Object.

#### Example

Activities that a stakeholder may wish to perform when accessing an email Information Object includes:

- Establish the email account from which a message originated.
- Establish the name of the person who sent the message.
- Establish the placement of the email message within an ongoing discussion.
- View the textual content of the message
- Establish the relational structure of text and images within the message
- View the visual appearance of the message as defined by the creator.
- Establish the route that an email took to reach the recipient.
- Add additional descriptive metadata or modify the subject line to enable the Object to be located at a later date.

The majority of the activities outlined above may be found in the expected behaviours. However, the stakeholder may have additional requirements that were not previously envisaged. The final bullet point in the example indicates that they require the ability to update the email with additional descriptive metadata. A behaviour may be distinct to a specific stakeholder. Other stakeholders, such as a curatorial institution, would not exhibit the behaviour.

---

<sup>12</sup> As noted in the introduction to FBS, Gero (1990) uses Behaviours to refer to the actions performed by an artefact, while Stalker (2002) refers to user behaviour. The latter interpretation is used here.

<sup>13</sup> Stalker (2002) highlights the widespread use of text messaging functionality as an actual behaviour that was not envisaged by early phone designers.

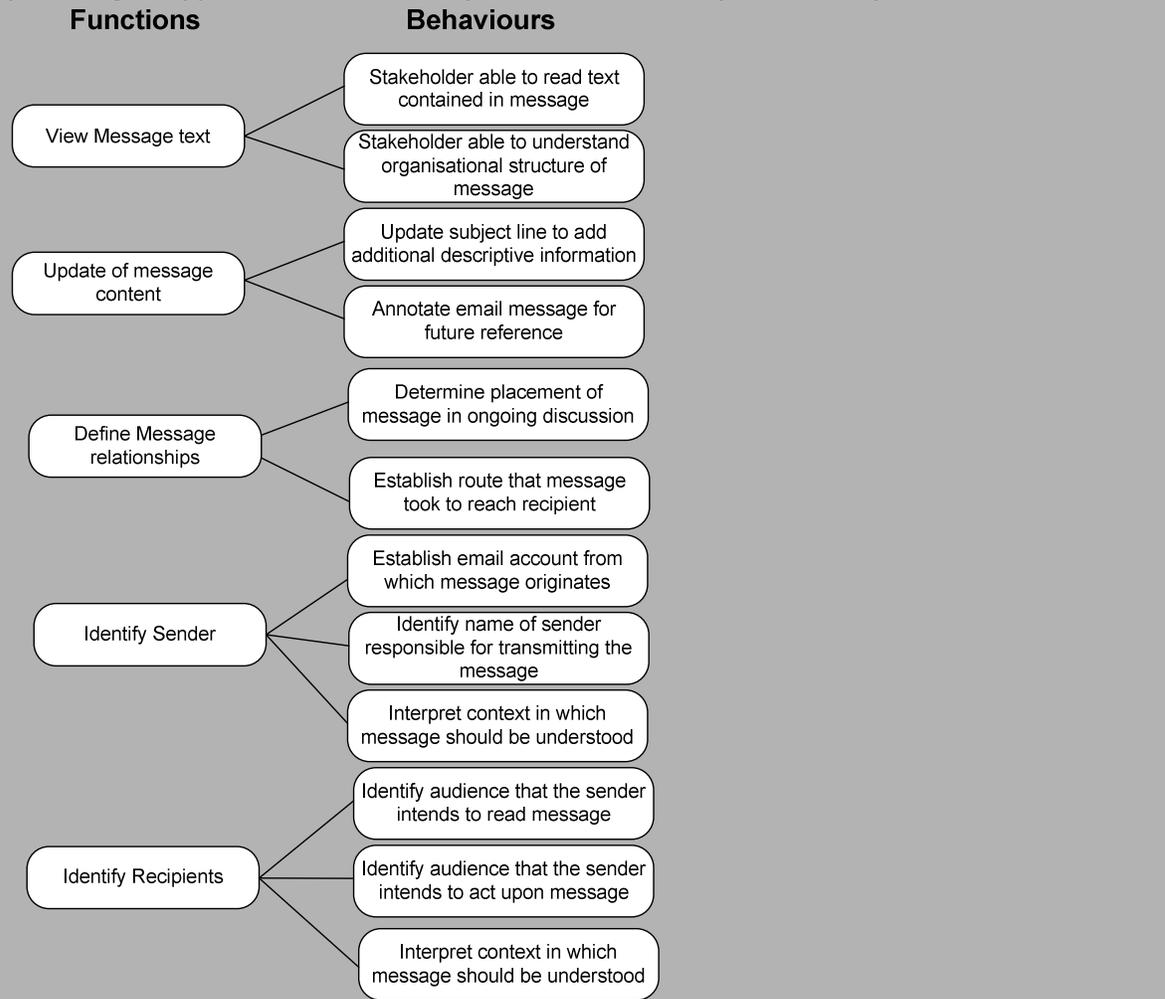
## 4. Classify behaviours into set of functions

The purpose of the fourth step is to classify the set of behaviours identified in the previous stage into a set of functions that subsequent manifestations of the Information Object should perform. A function refers to a specific design intention or purpose to be performed. In performing the activity, the evaluator may recognize that two or more behaviours should be associated with a single function. The evaluator should use the functional classification performed in step 5 of the object analysis as a classification guide. One or more new functions may need to be defined if the list of actual behaviours contains uses that were not previously recorded in the list of expected behaviours.

### Example

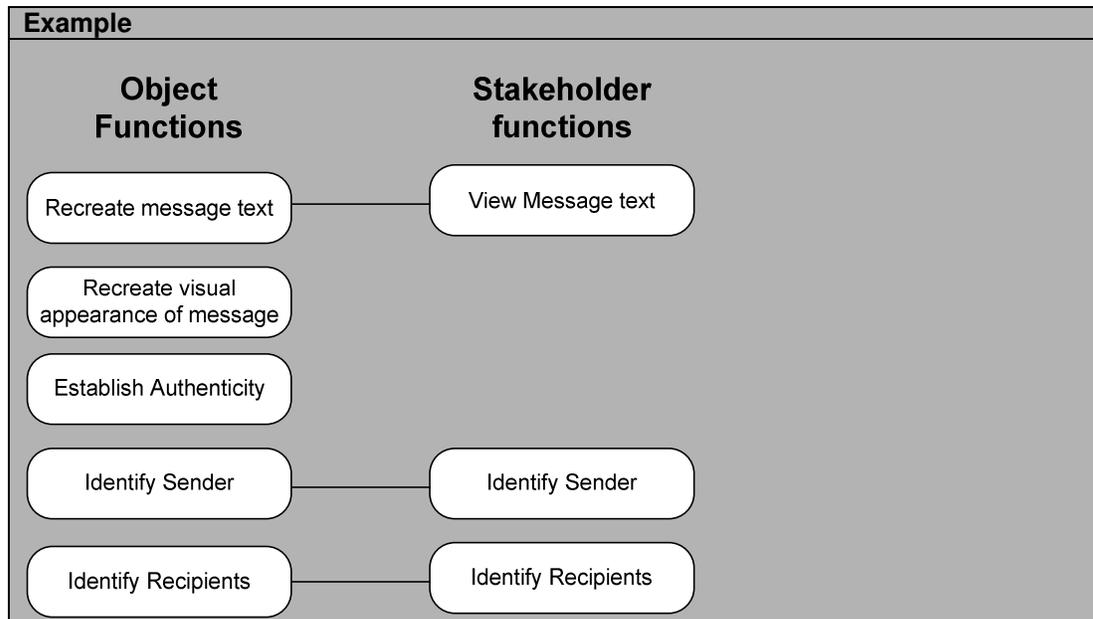
#### Email

In the example below, the stakeholder has expressed the need to perform many of the behaviours established in the list of expected behaviours. However, they have no interest in examining trace route information to establish authenticity. They have also expressed a requirement to be able to annotate the message content. The feature may not be present in the current manifestation of an object, but might be considered a worthwhile function to add when producing a copy of the Information Object in a new Data Object for use by the stakeholder.



## 5. Cross-match functions

The objective of the fifth step is to develop a list of the technical properties that are significant in performing the functions required by a stakeholder. To achieve the task, the set of functions identified for the stakeholder in step 4 should be cross-matched with the set of object type functions developed previously. The set of properties that are developed will enable the stakeholder to perform basic functions associated with the reproduction of the information content (e.g. view a still image, listen to an audio recording) and may include additional functions, as influenced by the type of activities that they wish to perform (e.g. verify authenticity).



## 6. Assign Acceptable value boundaries

The objective of the sixth step is to determine the value boundaries for properties that are acceptable to the stakeholder. The acceptable value boundaries may be used to assess the success of preservation action when creating an Information Object for use by a particular stakeholder. In some circumstances it may be impossible or impractical to transfer all properties when re-formulating an object. However, it may be questioned if the degree of accuracy is expected or required. Although the Information Object is not reproduced exactly, it may be sufficiently accurate to perform the functions required by a stakeholder. It may be feasible to assign quality thresholds for a minimum and maximum value for properties that are beneficial but not essential to the understanding of the Information Object's meaning or that allow some value variation without having a noticeable impact.

Four boundary constraints are currently recognized:

1. *Equality*: the property stored in the Record must be equal to one or more values stored in the metadata.
2. *Minimum*: if a numeric measurement is used, minimum indicates the lowest numeric value that is allowed. The minimum and maximum measurement types must be used in combination.
3. *Maximum*: if a numeric measurement is used, maximum indicates the highest number value that is allowed. For example, the highest sampling rate of an audio recording.
4. *Range*: the value is one of several that are recorded.

The type of values assigned to the acceptable boundary is likely to differ for each property. Examples may include numeric, text, or alphanumeric. Property value boundaries may also be set on the number of characters that are accepted.

An evaluator may take one of several approaches to obtaining information to populate the acceptable value boundary fields. They may observe the method in which the stakeholder uses the information and determine the quality level required to achieve a specific function. Alternatively, if the stakeholder is technically inclined, the evaluator may choose to explain the purpose of each property and ask them to consider the acceptable variation to perform the identified functions.

#### **Example**

##### *Sample rate acceptable boundary*

The sample rate of an audio recording has been measured as 48000 Hz. It may be found that the stakeholder is willing to accept a small degree of quality reduction, but does not wish it to be noticeable. In this example, the evaluator may assign a minimum boundary constrain of 44100 and a maximum of 96000.

##### Body text colour

The text colour of an email message has been identified as potentially significant for understanding the intrinsic meaning of the message. The stakeholder wishes to maintain the colour information, but it has been determined that it is sufficient to distinguish the text only and does not need to be an exact colour.

## **7. Review and finalise**

Finally, the evaluator should review the information gathered in the previous steps and consider if any revisions should be made. Pertinent questions to be raised at this stage may include:

1. Are there other behaviours that may be exhibited?
2. Would it be more appropriate to de-compose one behaviour into two sub-behaviours to provide a more accurate description of the activity?
3. Are there other functions that may be identified?

Once the evaluator is satisfied that they have completed the task, they may record the stakeholder functional requirements.

### 3.3 Reformulation

Reformulation in a design context refers to a process of re-developing an artefact to perform a revised set of functions or enable different behaviours. A digital object may be reformulated at several stages in a curation lifecycle. Prior to performing preservation action, a digital curator will (possibly unconsciously) select a subset of existing functions found in the source object that must be exhibited when it is transferred into a form for storage in an AIP. Similarly, when creating a DIP the object is likely to be re-formulated to perform functions required by users in the new Designated Community<sup>14</sup>.

In the FBS-based data model developed for the project, functions required in manifestations of the OAIS AIP and DIP are associated with one or more stakeholders. By adopting this approach, a list of technical properties may be developed that are significant in the context that the Information Object is used. For example, the functions established by Curatorial institution A may specify that properties associated with the header and message text are significant, while the functions specified for Curatorial institution B may take a more risk-adverse approach by specifying that all properties associated with the header, text and visual appearance of the message are significant.

To illustrate the re-formulation process, figure 7 indicates a simple workflow in which an Information Object created for use by Stakeholder A is re-formulated into an Information Object for use by stakeholder B. To perform the process, it is necessary to identify the set of high-level functions required by the stakeholder (as identified in the Stakeholder requirements analysis) and cross-match it with the Object properties necessary to perform each function. The re-formulation specification may be subsequently used to evaluate the success of the conversion.

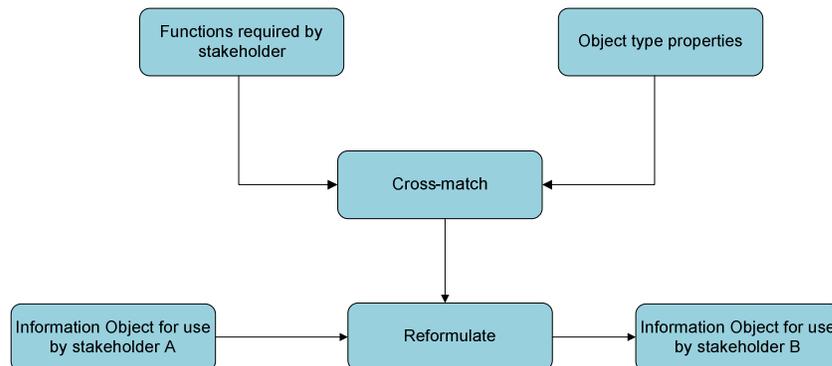


Figure 7: A simple re-formulation workflow

In a working environment, the workflow necessary to tailor an Information Object to the requirements of specific user types and validate that all required properties are present will be more complex. An abstract flowchart (figure 8) may be developed that indicate the set of activities performed to re-formulate an object (e.g. an email contained with a collection) using information captured in the Object Type and Stakeholder Requirements analysis.

The workflow begins when a user requests an object conforming to a specific object type/sub-type (e.g. an email, raster still image, audio recording, structured text document). If the user can

---

<sup>14</sup> For example, an Information Object created for a general audience may comprise a still image only, while an Information Object created for use by tutors and students may comprise a still image and Learning Object metadata that support its use in learning and teaching. Both manifestations of the Information Object will contain properties that are significant in the associated context.

be classified as one of the previously defined stakeholders (e.g. researcher, curator, tutor) a set of functions associated with the stakeholder may be obtained from an appropriate information source. If the user does not conform to an existing stakeholder classification or wishes to perform functions that cross-over two different stakeholders (e.g. genealogy) they may be given the option to select a set of behaviours associated with the object type<sup>15</sup>, which is used as a basis for defining a new set of functions.

In the second stage of the workflow, the requested object is analysed to ascertain that it can perform the set of functions requested by the user. One approach may involve the comparison of a set of properties associated with selected functions to metadata generated by a characterisation tool (e.g. JHOVE). Two responses may be produced as a result of the comparison:

1. The object contains all of the properties necessary to fulfill the specified functions and the requested object is re-formulated<sup>16</sup>. The re-formulated object must be subsequently validated to confirm that it contains all of the properties necessary to perform the required functions.
2. The object contains a subset or none of the required properties and, as a result, the user will be able to perform few or none of the specified functions. At this stage, the user may be asked if they wish to continue with the re-formulation activity or obtain the Object in its original form. If the former, the object is re-formulated to perform a subset of functions.

Finally, the re-formulated object is validated that it contains the requisite properties to perform each function. An invalid object may be deleted, at which point the workflow would return to an earlier stage. A valid object may be subsequently made available to the user for use.

---

<sup>15</sup> It is impractical to expect the user to manually define the set of actions that they wish to perform and associate them with an existing or new function in real-time – the process would be too laborious and may be technically difficult. However, the user may represent a new stakeholder type that the institution could analyse using the Stakeholder Requirements workflow at a later date.

<sup>16</sup> Additional actions may also be performed at the same time. For example, format conversion.

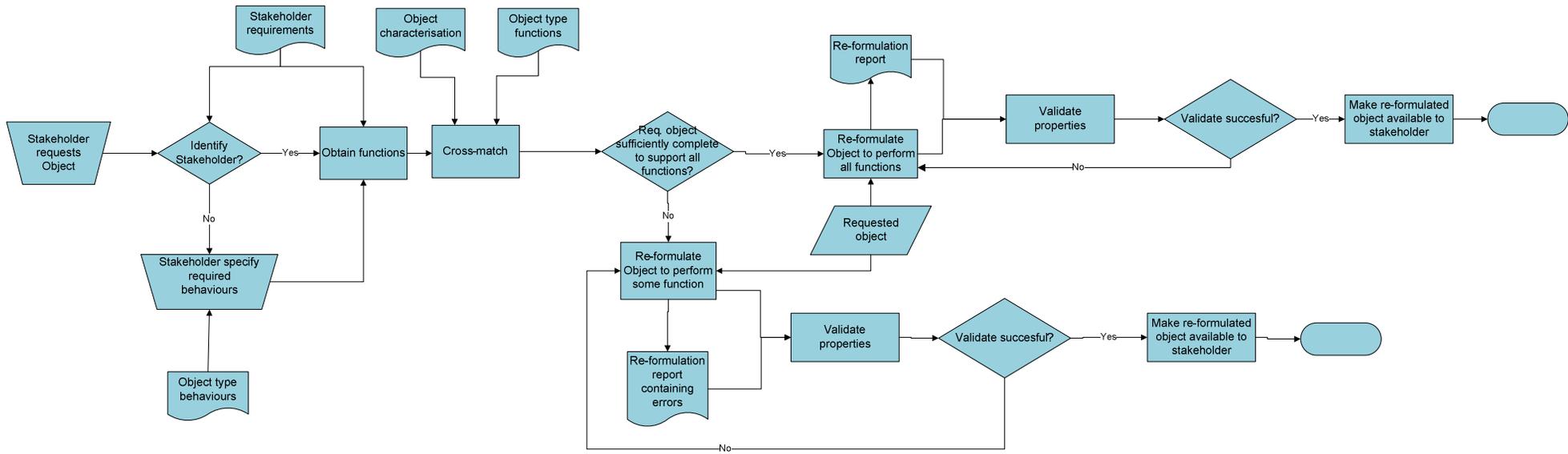


Figure 8: A flowchart indicating activities necessary to tailor an IO and establish the set of properties that are significant in a particular context

## ***Significant Properties Data Dictionary***

The InSPECT project also developed a data dictionary for recording the properties of an Information Object that are significant to a stakeholder. It was recognized that many institutions possess procedural lists and guidelines that indicate the properties associated with specific types of digital object. However, they are often stored in an unstructured form within an electronic document. The SP data dictionary supports the management of digital objects by enabling institutions to:

1. Record the components and properties of a digital object
2. Evaluate the subjective value of each component and property that represent the information content.
3. Assign quantitative and qualitative quality thresholds for the recreation of information content
4. Evaluate the recreation of the Information Object by comparing properties stored in a source and destination Data Object.
5. Obtain information regarding the ability to maintain each property when converting to a different encoding format, by querying a third-party service<sup>17</sup>.

The use of a data dictionary in its various implementations (e.g. XML, database) will enable an institution to store object type information and stakeholder requirements to be stored in the same environment as the digital object objects themselves.

The data model developed for use in the Significant Properties Data Dictionary adopts the Information Object approach specified in the OAIS RM and draws upon work undertaken by the PREMIS Working Group and National Archives Seamless Flow programme, among others to define a set of sub-units. In its final version developed for the InSPECT project, the data model defines four entities: an Information Object that represents a compound of many types of information (e.g. text, images, sounds, etc.) consisting of intellectual or technical components; Component that refers to a subset of the Information Object with which multiple properties are associated (e.g. a shape within a vector diagram; and Properties which represent the technical or semantic characteristics required to recreate the Object in part or in full. Each property will possess one or more values. Finally, an Agent must be associated with each entity to identify the stakeholder associated with the definition, classification and/or evaluate of a property within an Information Object. Future work on the Data Dictionary will re-develop the data model to better represent the FBS-derived methodology and is likely to draw upon related work being performed in the PLANETS project by Dappert & Farquhar (2008) and the University at Cologne (2008)

---

<sup>17</sup> The preservation services necessary to perform the task do not exist at the time of writing. Future work in this area may be performed by a technical registry, such as the Unified Digital Formats Registry or PRONOM.

## References

- Ashley, K. Davis, R & Pinsent, E. 2008. Significant Properties of e-Learning Objects (SPeLOs), v1.0.  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/2008sigprops.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx)
- Bearman, D. & Trant, J. 1998. Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process, *D-Lib Magazine*, June 1998.  
<http://www.dlib.org/dlib/june98/06bearman.html>
- Brown, A. 2008. White Paper: Representation Information Registries  
[http://www.planets-project.eu/docs/reports/Planets\\_PC3-D7\\_RepInformationRegistries.pdf](http://www.planets-project.eu/docs/reports/Planets_PC3-D7_RepInformationRegistries.pdf)
- CAMiLEON project. n.d. Creative Archiving at Michigan & Leeds: Emulating the Old on the New.  
<http://www.si.umich.edu/CAMiLEON/>
- CASPAR Project. n.d. Cultural, Artistic and Scientific Knowledge for Preservation, Access and Retrieval. <http://www.casparpreserves.eu/>
- Cedars Project. 2002. Cedars Guide To : Digital Collection Management.  
<http://www.leeds.ac.uk/cedars/guideto/collmanagement/>
- Coyne, M et al. 2007. The Significant Properties of Vector Images.  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/2008sigprops.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx)
- Coyne, M. & Stapleton, M. 2008. The Significant Properties of Moving Images.  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/2008sigprops.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx)
- Consultative Committee for Space Data Systems. 2002. Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-B-1, Blue Book  
<http://public.ccsds.org/publications/archive/650x0b1.pdf>
- Dappert, A. & Farquhar, A. 2009. Significance is in the Eye of the Stakeholder. ECDL 2009, LNCS 5714, pp 297-308.  
[http://www.planets-project.eu/docs/papers/Dappert\\_Significant\\_Characteristics\\_ECDL2009.pdf](http://www.planets-project.eu/docs/papers/Dappert_Significant_Characteristics_ECDL2009.pdf)
- Digital Curation Centre. 2008. The DCC Curation Life Cycle Model.  
<http://www.dcc.ac.uk/docs/publications/DCCLifecycle.pdf>
- Digital Preservation Testbed. 2003. From digital volatility to digital permanence: Preserving text documents. <http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=185>
- Dorst, K. & Vermaas, P.E. 2005, 'John Gero's Function-Behaviour-Structure Model of Designing: A Critical Analysis.', *Research In Engineering Design*, vol. 16, no. 1-2, pp. 17-26.  
<http://www.springerlink.com/content/v684602542mp8070/>
- Fallis, D. 2006. Social epistemology and information science. *ARIST* 40, pp 475-519.
- Gero J.S. 1990. Design Prototypes: A Knowledge Representation Schema for Design. *AI Magazine* 11(4): 26-36.  
<http://www.aaai.org/ojs/index.php/aimagazine/article/viewArticle/854>
- Heslop, H. Davis, S. & Wilson, A. 2002. An Approach to the Preservation of Digital Records  
[http://www.naa.gov.au/Images/An-approach-Green-Paper\\_tcm2-888.pdf](http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm2-888.pdf)
- Jones, S. Ross, S. & Ruusalepp, R. 2009. Data Audit Framework Methodology. Version 1.8  
<http://www.data-audit.eu/methodology.html>
- Knight, G. 2008a. Framework for the Definition of Significant Properties.  
<http://www.significantproperties.org.uk/outputs.html>

Project Acronym: InSPECT  
Version: 1.0  
Contact: Stephen Grace  
Date: 21 December 2009

Knight, G. 2008b.. Significant Properties Data Dictionary.  
<http://www.significantproperties.org.uk/outputs.html>

Knight, G. 2008c.. Deciding factors: Issues that influence decision-making on significant properties  
<http://www.significantproperties.org.uk/outputs.html>

Knight, G. 2009a. Significant Properties Testing Report: Digital Audio Recordings  
<http://www.significantproperties.org.uk/outputs.html>

Knight, G. 2009b., Significant Properties Testing Report: Electronic Mail  
<http://www.significantproperties.org.uk/outputs.html>

Lavoie, B.F. 2004. Technology Watch Report The Open Archival Information System Reference Model: Introductory Guide. [http://www.dpconline.org/docs/lavoie\\_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf)

MacNeil, H. et al. (2000). Authenticity Task Force Report (2000).  
[http://www.interpares.org/display\\_file.cfm?doc=ip1\\_atf\\_report.pdf](http://www.interpares.org/display_file.cfm?doc=ip1_atf_report.pdf)

Matthews, B. et al. 2008. The Significant Properties of Software: A Study, Version 5.7.  
[http://www.jisc.ac.uk/whatwedo/programmes/programme\\_preservation/2008sigprops.aspx](http://www.jisc.ac.uk/whatwedo/programmes/programme_preservation/2008sigprops.aspx)

Montague, L. 2009a. Significant Properties Testing Report: Raster Images  
<http://www.significantproperties.org.uk/outputs.html>

Montague, L. 2009b. Significant Properties Testing Report: Structured Text  
<http://www.significantproperties.org.uk/outputs.html>

PLANETS Project. n.d. Preservation and Long-term Access through NETworked Services.  
<http://www.planets-project.eu/>

PREMIS Editorial Committee (2008). PREMIS Data Dictionary for Preservation Metadata. Version 2.0. <http://www.loc.gov/standards/premis/>

Rauch, C. Strodl, S. & Rauber, A. 2005. Deliverable 6.4.1: A Framework for Documenting the Behaviour and Functionality of Digital Objects and Preservation Strategies.  
[http://www.dpc.delos.info/private/output/DELOS\\_WP6\\_d641\\_final\\_\\_vienna.pdf](http://www.dpc.delos.info/private/output/DELOS_WP6_d641_final__vienna.pdf)

RLG. 2002. Trusted Digital Repositories: Attributes and Responsibilities.  
<http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>

Rothenberg, J. & Bikson, 1999. T. Carrying Authentic, Understandable and Usable Digital Records Through Time: Report To the Dutch National Archives And Ministry of the Interior.  
[http://www.digitaleduurzaamheid.nl/bibliotheek/docs/final-report\\_4.pdf](http://www.digitaleduurzaamheid.nl/bibliotheek/docs/final-report_4.pdf)

Rusbridge, C. 2006. Excuse Me... Some Digital Preservation Fallacies?. Ariadne Issue 46.  
<http://www.ariadne.ac.uk/issue46/rusbridge/intro.html>

Stalker, R. 2002. A function-behaviour-structure framework for the lifecycle of an artefact.  
<http://www.ruthstalkerfirth.com/pdf/CC02.pdf>

Takeda, H. et al. n.d.. Analysis of Design Processes by Function, Behavior and Structure – Preliminary Reports.  
<http://www.kasm.nii.ac.jp/papers/takeda/95/DPW-paper.pdf>

Thibodeau, K. (2002). Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years  
<http://www.clir.org/pubs/reports/pub107/thibodeau.html>

University at Cologne (2008). Final XCDL Specification.  
[http://www.planets-project.eu/docs/reports/Planets\\_PC2-D7\\_FinalXCDLSpec\\_Ext.pdf](http://www.planets-project.eu/docs/reports/Planets_PC2-D7_FinalXCDLSpec_Ext.pdf)

Project Acronym: InSPECT  
Version: 1.0  
Contact: Stephen Grace  
Date: 21 December 2009

The National archives, n.d. Seamless Flow.  
[http://www.nationalarchives.gov.uk/electronicrecords/seamless\\_flow/default.htm](http://www.nationalarchives.gov.uk/electronicrecords/seamless_flow/default.htm)

Wilson, A. 2007. Significant Properties Report.  
[http://www.significantproperties.org.uk/documents/wp22\\_significant\\_properties.pdf](http://www.significantproperties.org.uk/documents/wp22_significant_properties.pdf)